

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2003, 04:28:29 ; Search time 3057.14 Seconds

(without alignments)

1140.147 Million cell updates/sec

Title: US-09-818-143-6

Perfect score: 1439

Sequence: 1 gggtacccctgttaacca.....agaaaaaaaaaaaaaa 1439

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

22781392 seqs, 12152238056 residues

1 number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST,*

1: em_estba:*

2: em_estchum:*

3: em_esttini:*

4: em_esttmr:*

5: em_estcov:*

6: em_estpli:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_eston:*

17: em_gss_hum:*

18: em_gss_invi:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gse_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1439	100.0	1439	10 AW88220	AW88220 MRRA3_Hum
2	1292	89.8	2310	11 BC03275	BC03275 Homo Sapiens
3	813.6	56.5	922	9 AL521507	AL521507 AL521507
4	804.6	55.9	922	9 AL521506	AL521506 AL521506

ALIGNMENTS

C	5	735.2	51.1	741	14 CN430710
C	6	706	49.1	738	10 BG748266
C	7	701.4	70.1	703	12 BQ019860
C	8	698.8	48.6	703	12 BM910547
C	9	698.5	73.1	703	14 CB81585
C	10	693.6	48.2	1064	13 BX35388
C	11	687	47.7	703	12 BG678659
C	12	677.4	47.1	733	12 BM914268
C	13	635.2	44.1	674	14 CA410929
C	14	626.6	43.5	668	12 BM701957
C	15	624.8	43.4	673	13 BQ418042
C	16	624.8	43.4	683	13 BU627928
C	17	607	42.2	760	13 BU170946
C	18	603.6	41.9	664	12 BM915562
C	19	596	41.4	766	14 CDD14758
C	20	590.4	41.0	636	13 BU729623
C	21	582.5	40.5	602	13 BU633800
C	22	574.4	39.9	594	14 CB052399
C	23	560.2	38.9	628	9 AA115295
C	24	555.6	38.6	574	14 CA777565
C	25	543.5	37.8	605	14 CA777311
C	26	541	37.6	587	14 CB36370
C	27	537.6	37.4	571	12 BM70342
C	28	537	37.3	571	12 BM030782
C	29	534.5	37.2	542	12 BQ019900
C	30	532	37.0	540	12 BM030553
C	31	527	36.6	535	9 AL043024
C	32	526	36.6	538	10 BP724489
C	33	525.8	36.5	642	12 BM61762
C	34	522	36.3	562	13 BU633007
C	35	517.4	36.0	530	9 AA625435
C	36	507	35.2	507	14 CAS15879
C	37	504	35.0	558	9 AI931849
C	38	499.8	34.7	649	10 BE966909
C	39	496.6	34.7	650	10 AW936987
C	40	496.2	34.5	501	9 AW70019
C	41	494	34.3	41	14 WA5522
C	42	483.2	33.6	516	9 AA873315
C	43	479.6	33.3	518	12 BM914490
C	44	476.2	33.1	501	9 AW059314
C	45	476	33.1	479	9 AI926527

RESULT 1

AW88220 LOCUS AW88220 DEFINITION 1439 bp mRNA linear EST 05-MAR-2001

DBDEFINITION MRRA3 Human matrix tissue expression library Homo Sapiens cDNA

Clone Incyte 1556751 similar to MRRA3 Matrix remodeling associated gene 3' mRNA sequence.

ACCESSION AW88220 KEYWORDS EST.

VERSION AW88220.1 GI:8050231

ORGANISM Homo sapiens (human)

SOURCE Homo Sapiens

REFERENCE 1. (bases 1 to 1439)

AUTHORS Walker,M.G. and Volkman,W.

TITLE Matrix-remodeling associated genes identified by co-expression

JOURNAL Unpublished

COMMENT Contact: Michael Walker

Incro Genomics

3160 Porter Drive, Palo Alto, CA 94304, USA

Email: mwalker@incro.com.

Fax: 650 845-5771

Location/Qualifiers 1..1439

/organism="Homo sapiens"

/mol_type="mRNA"

BASE COUNT		392 a	336 c	338 g	373 t
Query Match	100 %	Score 1439;	DB 10;	Length 1439;	
Best Local Similarity	100.0%	Pred. NO. 4e-208;			
Matches	1439;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	GAAGTATCCCTGTTAATCTACTTGTGGTTAAAGAGACCTTGCGTCACTGCCTCA	60		
Db	1	GAGTATCCCTGTTAATCTACTTGTGGTTAAAGAGACCTTGCGTCACTGCCTCA	60		
QY	61	TTCCTTGAAGATTAGCCCTGGCTCACTTCACTTATTCTCTCTGTCAGAA	120		
Db	61	TTCCTTGAAGATTAGCCCTGGCTCACTTCACTTATTCTCTCTGTCAGAA	120		
QY	121	AGAGAGAGAGAGAGAACATTACCGAAACCCCTCCCTCCAGAGAGAA	180		
Db	121	AGAGAGAGAGAGAGAACATTACCGAAACCCCTCCCTCCAGAGAGAA	180		
QY	181	CAATATGTAATTCTCTAGAAACCTCAGAGCTGAGAGCAGAAAGAA	240		
Db	181	CAATATGTAATTCTCTAGAAACCTCAGAGCTGAGAGCAGAAAGAA	240		
QY	241	CTCAGGCGGGCTCTCGTCCCGCCCGAGCCAGACTATGTCAGAGGCGTCT	300		
Db	241	CTCAGGCGGGCTCTCGTCCCGCCCGAGCCAGACTATGTCAGAGGCGTCT	300		
QY	301	GTTCCTCTCCGGAGACTCCAGAAATTCTCTCAAGGAGAAACAGGGATGCGT	360		
Db	301	GTTCCTCTCCGGAGACTCCAGAAATTCTCTCAAGGAGAAACAGGGATGCGT	360		
QY	361	GTTCGCCAAAGCAGCGCGCTCCAAACCCATGGTAGATTAAGATGGCC	420		
Db	361	GTTCGCCAAAGCAGCGCGCTCCAAACCCATGGTAGATTAAGATGGCC	420		
QY	481	GCCAGCTGGCTTTAACAACTTAGAGACTTCCGGAGCTGCCCTGGAA	540		
Db	481	GCCAGCTGGCTTTAACAACTTAGAGACTTCCGGAGCTGCCCTGGAA	540		
QY	541	GRCCCCAGCTCCAGGAACTTAAGAGACGGCTTGCGAAACACATG	600		
Db	421	GRCCCCAGCTCCAGGAACTTAAGAGACGGCTTGCGAAACACATG	480		
QY	541	GGAAACGGGCTGGCAAGACACTCAAGTTCAGCTTGCGCTGGTT	600		
Db	541	GGAAACGGGCTGGCAAGACACTCAAGTTCAGCTTGCGCTGGTT	600		
QY	601	GAATCTCACATCATCTGCTCTGGCCCTTCTCTGGTGGAAAGGGAG	660		
Db	601	GAATCTCACATCATCTGCTCTGGCCCTTCTCTGGTGGAAAGGGAG	660		
QY	661	GTTCAGCGCTCTAAATGCTGTTGGTCAAGGGTGGCTCAACCAATA	720		
Db	661	GTTCAGCGCTCTAAATGCTGTTGGTCAAGGGTGGCTCAACCAATA	720		
QY	721	GTCGCTTCTGAGGCCAGTGAACCTCAAGGCTTAAGCCCAATTCTAG	780		
Db	721	GTCGCTTCTGAGGCCAGTGAACCTCAAGGCTTAAGCCCAATTCTAG	780		
QY	781	AGAAAGCGCTGAGCTCCAGTTCATTAAGAGCCAAAGCTGGTCTCTCA	840		
Db	781	AGAAAGCGCTGAGCTCCAGTTCATTAAGAGCCAAAGCTGGTCTCTCA	840		
QY	841	GAAGACTAGTTCTGCTGTTAGTACAGCTGCTGTTGAGTTAGAAGGA	900		
Db	841	GAAGACTAGTTCTGCTGTTAGTACAGCTGCTGTTGAGTTAGAAGGA	900		
QY	901	GTACGAAACACAAATAGCGCTGCTGCTGAGGAGCTGTTAGAAGGA	960		
Db	901	GTACGAAACACAAATAGCGCTGCTGAGGAGCTGTTAGAAGGA	960		
QY	961	TGGATGGATGGATGATGATGATGATGATGATGATGATGATGATGAGAA	1020		
Db	961	TGGATGGATGGATGATGATGATGATGATGATGATGATGATGAGAA	1020		
QY	1021	ACAAACUAAGCTGATGGACATTATTGCGGCTGAGGGAGGGCGC	1140		
Db	1021	ACAAACUAAGCTGATGGACATTATTGCGGCTGAGGGAGGGCGC	1140		
QY	1141	AGCCAGTGATTATTAGCTCTAGAGCAGACTCTGCCAGGATGCTGAC	1200		
Db	1141	AGCCAGTGATTATTAGCTCTAGAGCAGACTCTGCCAGGATGCTGAC	1200		
QY	1201	CTGAGATGGTCCACGCTCTACACTCTCTCCCGTGGATACTGGACTGT	1260		
Db	1201	CTGAGATGGTCCACGCTCTACACTCTCTCCCGTGGATACTGGACTGT	1260		
QY	1261	CCTGGCCGCAATTATTGAGCTGAGGAACTCTTGGGATGCTGATGAC	1320		
Db	1261	CCTGGCCGCAATTATTGAGCTGAGGAACTCTTGGGATGCTGATGAC	1320		
QY	1321	GGTAGCTTCAATAAGCGAGTACATTAAGTGAATGATCTTCTGATGTT	1380		
Db	1321	GGTAGCTTCAATAAGCGAGTACATTAAGTGAATGATCTTCTGATGTT	1380		
QY	1381	TTAGAAGTAAACAATTAAAGCTACCAATTATGAGAAAAAA	1439		
Db	1381	TTAGAAGTAAACAATTAAAGCTACCAATTATGAGAAAAAA	1439		
RESULT 2					
BC032745					
LOCUS					
DEFINITION					
Hom sapiens, DKFZP43N161 protein, clone IMAGE:5574690, mRNA.					
VERSION					
BC032745.1					
KEYWORDS					
Homo sapiens (human)					
SOURCE					
Organism					
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Butherida; Primates; Catarrhini; Hominidae; Homo.					
1 (bases 1 to 2310)					
AUTHORS					
Straubhaar, R.					
DIRECT SUBMISSION					
Submitted (06-JUN-2002) National Institutes of Health, Mammalian Gene Collection (MGCI), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
NIH-MCG Project URL: http://mgcc.nci.nih.gov					
REMARK					
COMMENT					
Contact: MGCC help desk					
Email: cabpb-help@mail.nih.gov					
Tissue Procurement: NTC					
DNA Library Preparation: Life Technologies, Inc.					
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NSIC), Gaithersburg, Maryland,					
Web site: http://www.nisc.nih.gov/					
Contact: nisc.mc@nih.nih.gov					
Akhter, N.; Ayele, K.; Beckstrom-Sternberg, S.M.; Benjamin, B.; Blakesley, R.W.; Bouffard, G.G.; Green, K.; Brinkley, C.; Brooks, S.; Dietrich, N.L.; Granite, S.; Guan, X.; Gupta, J.; Hashiguchi, S.; Hansen, N.; Ho, L.; Karitis, E.; Kwong, P.; Iaric, P.; Legepi, R.; Maduro, O.L.; Maslloiu, B.; Maskeri, B.; Matsumoto, S.D.; McCloskey, J.C.; McDowell, J.; Pearson, R.; Stancropop, S.; Thomas, J.D.; Touchman, J.W.; Turgeon, C.; Vogt, J.H.; Walker, M.A.; Weatherby, K.D.; Wiggins, L.; Young, A.; Zhang, L.; and Green, E.D.					

GenCore...version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on:

August 13, 2003, 03:56:05 ; Search time 86,2656 Seconds

Title: perfect score: US-09-818-143-6

Sequence: 1 gagtatccatgtttatca.....aaaaaaa 1439

Scoring table: IDENTITY_NUC Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

1 number of hits satisfying chosen parameters:

1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NU:

1: /cggn2_6/prodata/2/ina/5A_COMB.seq; *
2: /cggn2_6/prodata/2/ina/5B_COMB.seq; *
3: /cggn2_6/prodata/2/ina/6A_COMB.seq; *
4: /cggn2_6/prodata/2/ina/6B_COMB.seq; *
5: /cggn2_6/prodata/2/ina/PCITS_COMB.seq; *
6: /cggn2_6/prodata/2/ina/batchfileseq.seq; *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1.64	87.8	4772	4 US-09-424-970B-133
2	4.2	218	4	US-08-423-463-14
3	45.4	3.2	374	2 US-08-322-768A-35
4	44.4	3.1	15331	3 US-09-118-155-16
5	4.4	17337	3	US-09-128-155-17
6	44.2	3.1	1315	4 US-09-721-822A-10
7	43.8	3.0	367	2 US-08-322-768A-9
8	43.6	3.0	487	3 US-08-420-00-1
9	43.6	3.0	1172	1 US-07-945-288-9
10	43.6	3.0	1172	1 US-07-942-821-9
11	43.6	3.0	1172	1 US-08-451-809-9
12	43.6	3.0	1172	1 US-08-451-411-9
13	43.6	3.0	1172	5 PCT-US91-05251-9
14	43.4	3.0	7218	1 US-08-222-463-14
15	43	3.0	3693	3 US-09-056-226-1
16	42.6	3.0	3207	1 US-07-946-497-1
17	42.6	3.0	3207	1 US-08-463-322-1
18	42.6	2.8	3207	2 US-08-478-882-1
19	42	2.9	11558	5 PCT-US91-06251-3
20	41.8	2.9	16389	4 US-09-711-154-3
21	40.6	2.8	2089	4 US-09-69-417-19
22	40	2.8	32654	4 US-09-801-191A-3
23	39.6	3	450	3 US-09-395-902-155
24	39.6	2.8	3347	4 US-09-105-318
25	39.6	2.8	3347	4 US-09-136-457-318
26	39.2	2.7	957	4 US-09-128-972-12
27	39.2	2.7	1091	4 US-09-328-965-1

ALIGNMENTS

28	39.2	2.7	1474	3	US-08-821-994-54	Sequence 64, AppI
29	39.2	2.7	1493	6	5340934-5	Patent No. 5340934
30	39.2	2.7	3850	3	US-08-223-799-34	Sequence 34, AppI
31	38.8	2.7	855	3	US-08-066-619-30	Sequence 130, App
32	38.8	2.7	855	3	US-08-906-616-130	Sequence 130, App
33	38.8	2.7	855	3	US-09-012-431-330	Sequence 130, App
34	38.8	2.7	855	3	US-09-012-692-330	Sequence 130, App
35	38.8	2.7	855	3	US-08-906-613-330	Sequence 130, App
36	38.8	2.7	1190	4	US-08-390-207-1	Sequence 1, AppI
37	38.8	2.7	1912	2	US-08-487-828B-13	Sequence 13, AppI
38	38.6	2.7	7859	1	US-07-854-596B-4	Sequence 4, AppI
39	38.6	2.7	7859	2	US-08-450-905B-15	Sequence 15, AppI
40	38.6	2.7	7859	3	US-08-982-75P-15	Sequence 15, AppI
41	38.6	2.7	1378	4	US-09-149-476-208	Sequence 208, App
42	38.4	2.7	2323	4	US-09-149-476-208	Sequence 24, AppI
43	38.4	2.7	4072	3	US-09-272-496-7	Sequence 7, AppI
44	38.2	2.7	711	3	US-08-998-416-786	Sequence 786, App

PA	(INCYT) INCYTE PHARM INC.
XX	PT Walker MG, Volkman W, Klingler TM.
DR	WPI; 2000-317934/27.
DR	P-PSDB; AAY70744.
XX	PT protein co-expressed with matrix-remodeling proteins, useful in the diagnosis and treatment of cancer, cardiomyopathy, arthritis, angiogenesis, diabetic necrosis, atherosclerosis, fibrosis, and ulceration.
XX	PT Claim 1; Page 40-41; 55pp; English.
CC	The present sequence is NBSQ gene that is co-expressed with one or more known matrix-remodeling genes in a number of biological samples using an expression vector. This sequence was identified from the incyte clone 156751. The gene, protein, and antibody sequences can be used in the diagnosis and treatment or prevention of a disease associated with its altered expression. The diseases that can be treated are matrix-remodeling diseases, including cancer, cardiomyopathy, arthritis, angiogenesis, diabetic necrosis, atherosclerosis, fibrosis, and ulceration.
CC	Sequence 1439 BP; 392 A; 336 C; 338 G; 373 T; 0 other;
XX	Query Match Best Local Similarity 100.0%; Score 1439; DB 21; Length 1439; Matches 1439; Conservative 100.0%; Pred. No.; 0 other;
QY	1 GAGATGCCCTGTTAAATCACTTGTGGTTAAGAGAACCTTGCTGTCCTTA 60
Db	1 GAGATCCCTGTTAAATCACTTGTGGTTAAGAGAACCTTGCTGTCCTTA 60
QY	61 TTCTTGTAGAGTTAGCCCTGGCTCACATTCACTTGTGGCTCAGAA 120
Db	61 TTCTTGTAGAGTTAGCCCTGGCTCACATTCACTTGTGGCTCAGAA 120
QY	121 AGAAAAGAAAAGAACAAATTACCGAAACCCCTCCCTCCACAGAGGAGCTTG 180
Db	121 AGAAAAGAAAAGAACAAATTACCGAAACCCCTCCCTCCACAGAGGAGCTTG 180
QY	121 AGAAAAGAAAAGAACAAATTACCGAAACCCCTCCCTCCACAGAGGAGCTTG 180
Db	121 AGAAAAGAAAAGAACAAATTACCGAAACCCCTCCCTCCACAGAGGAGCTTG 180
QY	181 CAATGTTAATTCTCTAGAAATCTCTAGAGCTGAGGAAAGAAATCTGGCT 240
Db	181 CAATGTTAATTCTCTAGAAATCTCTAGAGCTGAGGAAAGAAATCTGGCT 240
QY	241 CTAGGGGGGCTCTGGCTCCCGCAGGCCAGACTATGGTCAAGGGCGCTT 300
Db	241 CTAGGGGGGCTCTGGCTCCCGCAGGCCAGACTATGGTCAAGGGCGCTT 300
QY	301 GTRCTCTCCGGGAACTCAGAAATTCTCTCAAGGAGAAACGGGCTGCTT 360
Db	301 GTRCTCTCCGGGAACTCAGAAATTCTCTCAAGGAGAAACGGGCTGCTT 360
QY	361 GTGGCAAAGCGCGCGCGCCGCAAAACCCCAGTGTGTGATGATGGCC 420
Db	361 GTGGCAAAGCGCGCGCGCCGCAAAACCCCAGTGTGTGATGATGGCC 420
QY	421 GRCGCCGAGCCCTCCAGCAACTTAAGAGACGGCTTGTGAACCAATGG 480
Db	421 GRCGCCGAGCCCTCCAGCAACTTAAGAGACGGCTTGTGAACCAATGG 480
QY	481 GCGCGCTGGGCTTAACTGAGACTTCCGAGTGGCTGGAGAGCTGGC 540
Db	481 GCGCGCTGGGCTTAACTGAGACTTCCGAGTGGCTGGAGAGCTGGC 540
QY	541 GAAACCGGGCTTGCGAGAACACTGAGTTCTCTGGCTTGTGAACCAATGG 600
Db	541 GAAACCGGGCTTGCGAGAACACTGAGTTCTCTGGCTTGTGAACCAATGG 600
QY	601 GATCTCCACATCATGCTTCTCTGGCTTGTGAACCAATGGAGGAGG 660
Db	601 GATCTCCACATCATGCTTCTCTGGCTTGTGAACCAATGGAGGAGG 660
QY	661 GTTCAGGCCTAAAGAGCTTGTGATCAGGAGTGGCTCCAAACCAAATAGGCC 722
Db	661 GTTCAGGCCTAAAGAGCTTGTGATCAGGAGTGGCTCCAAACCAAATAGGCC 722
QY	721 GTGCCCTTCGAGGCCGTGAGCTCAGCTCCAGGTTAAAGCAGCTTCAGCAG 78
Db	721 GTGCCCTTCGAGGCCGTGAGCTCAGCTCCAGGTTAAAGCAGCTTCAGCAG 78
QY	781 AGAACGCTGAGAGCTGAGCTGAGCTGAGCTGGCTCTCTGCTTA 84
Db	781 AGAACGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGGCTCTCTGCTTA 84
QY	841 GAAAGTAGCTTGTGGCTGATGAGCTGAGCTGGCTCTCTGATTTAGGAGAT 90
Db	841 GAAAGTAGCTTGTGGCTGATGAGCTGAGCTGGCTCTCTGATTTAGGAGAT 90
QY	901 GTACGAAACAAATGGGTGCTGTCGTCAGCAAGAGGAGACAGGATGCTGAG 96
Db	901 GTACGAAACAAATGGGTGCTGTCGTCAGCAAGAGGAGACAGGATGCTGAG 96
QY	961 TGGATGAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 10
Db	961 TGGATGAGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 10
QY	1021 ACACACAAAGTGTGGAAACATTATGTTGGTGTGGCTGAGGGAAAGTCGAGC 10
Db	1021 ACACACAAAGTGTGGAAACATTATGTTGGTGTGGCTGAGGGAAAGTCGAGC 10
QY	1021 ACACACAAAGTGTGGAAACATTATGTTGGTGTGGCTGAGGGAAAGTCGAGC 10
Db	1021 ACACACAAAGTGTGGAAACATTATGTTGGTGTGGCTGAGGGAAAGTCGAGC 10
QY	1081 TTGGGACCTTGTAGAGCGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 11
Db	1081 TTGGGACCTTGTAGAGCGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 11
QY	1141 AGCCAGTCATTATTTAACCTTGTAGAGCAACTCTCTGCGGAAATGCTGACC 12
Db	1141 AGCCAGTCATTATTTAACCTTGTAGAGCAACTCTCTGCGGAAATGCTGACC 12
QY	1201 CTGAGATGGTCACGATCTCTACGCTCTCTCTCGTGGATACTGACTGCTG 12
Db	1201 CTGAGATGGTCACGATCTCTACGCTCTCTCGTGGATACTGACTGCTG 12
QY	1261 CCTCTGCCCATCTCTCAGGATATCCATGAGCTTAATTCTACTCTGTAC 13
Db	1261 CCTCTGCCCATCTCTCAGGATATCCATGAGCTTAATTCTACTCTGTAC 13
QY	1321 GGATGCTCATAAGCCGACTATGACTTAAATGAGTATCTTGTGATCCTT 13
Db	1321 GGATGCTCATAAGCCGACTATGACTTAAATGAGTATCTTGTGATCCTT 13
QY	1381 TTAGAGGTAACAAATTAAAGCTCAATATGAGAAAAAAAGAAAAA 1439
Db	1381 TTAGAGGTAACAAATTAAAGCTCAATATGAGAAAAAAAGAAAAA 1439
QY	RESULT 2
Db	AAL42450
QY	ID AAL42450 Standard; DNA; 1439 BP.
Db	XID AAL42450
AC	AAL42450;
XX	DT 11-JUL-2002 (first entry)
XX	DE Human matrix-remodeling-associated nucleotide 6.
XX	Human; ds; matrix-remodeling gene; extracellular matrix; gene; matrix-remodeling-associated nucleotide; screening; matrix remodeling-associated disease; angiogenesis; arthritis; atherosclerosis; cancer; cardiomyopathy; diabetic necrosis; fibrosis; ulceration.
OS	Homo sapiens.
XX	KEV Coronation/Quintessential

FT PT
 FT PT
 XX PN
 XX US200219000-A1.
 XX /tags: a
 XX /product: "Human matrix-remodeling-associated protein 2"
 XX
 PR 09-OCT-1999; 98US-0169289.
 XX
 PT (WALK/) WALKER M G.
 PA (VOLK/) VOLKUTH W.
 PA (KLIN/) KLINGLER T M.
 XX
 PT Walker MG, Volkuth W, Klingler TM,
 XX DR
 WPI; 2002-338319/37.
 P-PSDB; AAO14785.
 PR
 XX
 PT
 CC associated with matrix-remodeling such as angiogenesis, arthritis and
 CC cancer -
 PT
 XX
 PS Claim 1; Page 21; 63pp; English.
 XX
 CC The invention comprises human nucleotide sequences which are co-expressed
 CC with matrix-remodeling genes. Matrix-remodeling is associated with the
 CC construction, destruction and reorganisation of extracellular matrix
 CC components. The matrix-remodeling-associated nucleotides of the invention
 CC are useful for screening for and purifying ligands that specifically bind
 CC to the nucleotides of the invention. The matrix-remodeling-associated
 CC nucleotides of the invention are also useful in the diagnosis, prognosis
 CC prevention, treatment and evaluation of therapies for diseases associated
 CC with matrix remodelling (e.g. angiogenesis, arthritis, atherosclerosis,
 CC cancer, cardiomyopathy, diabetic necrosis, fibrosis and ulceration). The
 CC present DNA sequence represents a human matrix-remodeling-associated
 CC nucleotide of the invention.
 XX
 SQ Sequence 1439 BP; 392 A; 336 C; 338 G; 373 T; 0 other;
 Query Match 100.0%; Score 1439; DB 24; Length 1439;
 Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 1439; Conservative 0;
 PR
 1 GAGTATCCCTGTTAACCTTGCGTTAAGAGCCTTGATCTGCCTCA 60
 1 GAGTATCCCTGTTAACCTTGCGTTAAGAGCCTTGATCTGCCTCA 60
 OY 61 TTCTTGAGGTTCGGCTGCTGCTACCTTCTCTCTCTCTCAAGAA 12
 DB 61 TTCTTGAGGTTCGGCTGCTGCTACCTTCTCTCTCTCAAGAA 12
 121 AGAGAGAAAAGAGACAAATTACCGAGAACCCCTCCCTCCCATGAGAGA 12
 OY 121 AGAGAGAAAAGAGACAAATTACCGAGAACCCCTCCCTCCCATGAGAGA 12
 DB 181 CAAATGTTAACCTTGCGTTAACCTTGCGTTAACCTTGCGTTAACCTTGCGTT 24
 OY 181 CAAATGTTAACCTTGCGTTAACCTTGCGTTAACCTTGCGTTAACCTTGCGTT 24
 DB 241 CTGAGGTGCGCTCTCGCCGCGCCAGGGCGCT 30
 OY 241 CTGAGGTGCGCTCTCGCCGCGCCAGGGCGCT 30
 DB 301 GTTCCTCCCGGACTCGAGAATTCTCTCTCAAGAGAGAACGGCATGCC 36
 OY 301 GTTCCTCCCGGACTCGAGAATTCTCTCTCAAGAGAGAACGGCATGCC 36
 DB 361 GTTGGCAAACGGCGGGCTCCAAACCCATGTTGATTAAGTGTGGCC 42

Db	3.61	GTTGGCAAACCGCGCGCGCTCCAAACCCATGTGTAAGTGTGCC	4.20
QY	4.21	GTCGCCASGCTCCCGCGAATTAAAGAGCAGGACTTGTGAAACCAACATG	4.80
Db	4.21	GTCGCCAGGCTCCAGGCCAACATTAAAGAGCAGGACTTGTGAAACCAACATG	4.80
QY	4.81	GCCAGCTGGCTTTAACCTAGAGACTTCCGGAGCTGCCTGCGACAGCGCTCG	5.40
Db	4.81	GCCAGCTGGCTTTAACCTAGAGACTTCCGGAGCTGCCTGCGACAGCGCTCG	5.40
QY	5.41	GGAAACGGGTTGCCAGAGACACTCAGTTCTCATGCCCTTGGTCCCCAA	6.00
Db	5.41	GGAAACGGGTTGCCAGAGACACTCAGTTCTCATGCCCTTGGTCCCCAA	6.00
QY	6.01	GAATTCACATCATTGTTCTGCGCTTCTTGAGGAGCTGCCTCAAACCAATTACGGC	7.20
Db	6.01	GAATTCACATCATTGTTCTGCGCTTCTTGAGGAGCTGCCTCAAACCAATTACGGC	7.20
QY	6.61	GTTCGAAGCTTAAMATGTGCTTGTAGGAGCTGCGCTTAAAGCCAAATTACGGC	7.20
Db	6.61	GTTCGAAGCTTAAMATGTGCTTGTAGGAGCTGCGCTTAAAGCCAAATTACGGC	7.20
QY	7.21	GTCGCCCTTGAGGCACTGTAGCTCGCCCTAACGGTTAAAGCCACATTGAG	7.80
Db	7.21	GTCGCCCTTGAGGCACTGTAGCTCGCCCTAACGGTTAAAGCCACATTGAG	7.80
QY	7.81	AGAAAGCTAGAGCTCGCAGGTTCTTAAAGAGGCGAAAGCACTGTTCTCTTA	8.40
Db	7.81	AGAAAGCTAGAGCTCGCAGGTTCTTAAAGAGGCGAAAGCACTGTTCTCTTA	8.40
QY	8.41	GAAAGTAGGTTCTGGCTGAGTGAAGCTGCGCTCTTGATTTTGTGAAGGAT	9.00
Db	8.41	GAAAGTAGGTTCTGGCTGAGTGAAGCTGCGCTCTTGATTTTGTGAAGGAT	9.00
QY	9.01	GTACCTAAACAAATAGGCTTGCGCTGCAAGAGACAGCAGGATGATGATA	9.60
Db	9.01	GTACCTAAACAAATAGGCTTGCGCTGCAAGAGACAGCAGGATGATGATA	9.60
QY	9.61	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATA	10.20
Db	9.61	TGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATA	10.20
QY	10.21	ACAAACCAAGCTGATGGAAACAATTAAATGCGGGCTGTGAGGGGAAAGGGCG	10.80
Db	10.21	ACAAACCAAGCTGATGGAAACAATTAAATGCGGGCTGTGAGGGGAAAGGGCG	10.80
QY	10.81	TTTGGCAGTTTGAGGAGGGTACAGAGCTCTGCGCTGCTGGTGTGAGGGGAAAGTCGGCG	11.40
Db	10.81	TTTGGCAGTTTGAGGAGGGTACAGAGCTCTGCGCTGCTGGTGTGAGGGGAAAGTCGGCG	11.40
QY	11.41	AGCCAGTCATTATTAAAGCTCTGAGAACAACTTAATGCGGGAAAGTCGGCG	12.00
Db	11.41	AGCCAGTCATTATTAAAGCTCTGAGAACAACTTAATGCGGGAAAGTCGGCG	12.00
QY	12.01	CTGAGCTGCTCCACGCMCTCTACACTCTCTCTCGTGGGAAACTGCTG	12.60
Db	12.01	CTGAGCTGCTCCACGCMCTCTACACTCTCTCGTGGGAAACTGCTG	12.60
QY	12.61	CCTCTGGGCCATCTCTCTACGGCATCTCATGAGCTTAACTCTCTGATAC	13.20
Db	12.61	CCTCTGGGCCATCTCTCTACGGCATCTCATGAGCTTAACTCTCTGATAC	13.20
QY	13.21	GGTAGCTCCATAAGCGAGTAACCTAAATGAGGATCTTGTAACTGTTT	13.80
Db	13.21	GGTAGCTCCATAAGCGAGTAACCTAAATGAGGATCTTGTAACTGTTT	13.80
QY	13.81	TTAGAGCTAAACATTAAATGCTACAAATACTGAGAAAAGAAAAGAAA	14.39
Db	13.81	TTAGAGCTAAACATTAAATGCTACAAATACTGAGAAAAGAAAAGAAA	14.39

AC	ABX77619	standard; cDNA; 1439 BP.
XX	09-APR-2003	(first entry)
XX	DT	
XX	DB	differentially expressed breast cancer associated cDNA #114.
XX	KW	Breast cancer; differential gene expression; BC-cDNA;
XX	KW	breast cancer diagnosis; breast cancer monitoring;
XX	KW	breast cancer treatment; breast cancer staging; gene; ss.
OS	Hom sapiens.	
PN	US2002156263-A1.	
XX	PD	24-OCT-2002.
XX	04-OCT-2001	2001US-0974298.
PS	WPI	2003-182653/18.
PA	(CHEN/)	CHEN H.
PA	PI	Chen H.
XX	XX	New cDNAs, which are differentially expressed in (metastatic) breast cancer useful for diagnosing or staging, breast cancer, or for monitoring the treatment of breast cancer in an individual.
XX	XX	Claim 1; SEQ ID NO 148; 30pp; English.
XX	XX	The invention describes a combination of cDNAs (designated BC-cDNAs), which are differentially expressed in breast cancer. The combination includes 152 cDNA sequences, or their complements. The protein encoded by any of these BC-cDNAs is useful for screening several molecules or compounds to identify at least one ligand that specifically binds the protein, producing or preparing polyclonal or monoclonal antibodies, or purifying antibodies from a sample. The antibodies, which specifically bind the protein, differentially expressed in breast cancer, is useful for detecting the expression of a protein in a sample. The BC-cDNAs are also useful for diagnosing, monitoring the treatment of, or staging, breast cancer. This sequence represents a differentially expressed
CC	Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from the US Patent office at seqdata.uspto.gov/Sequence.html?seqID=20020156263.	
XX	XX	Sequence 1439 BP; 392 A; 336 C; 338 G; 373 T; 0 other;
XX	XX	Query Match Local Similarity 100.0%; Score 1439; DB 25; Length 1439; Matches 1439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1	1	GAGATCCTGTTAACACTTGTAAAGAGGCCCTGGTCAGCTCA
2	1	GAGATCCTGTTAACACTTGTAAAGAGGCCCTGGTCAGCTCA
3	1	GAGATCCTGTTAACACTTGTAAAGAGGCCCTGGTCAGCTCA
4	61	TTCCTTGAGAGTTACCCCTGGCTACTTCTCTTCTCAGAA
5	61	TTCCTTGAGAGTTACCCCTGGCTACTTCTCTCAGAA
6	121	AGAAGAAAGAAAGAAATTACCCAGAAACCTCTCCGAGCTG
7	121	AGAAGAAAGAAAGAAATTACCCAGAAACCTCTCCGAGCTG
8	181	CAATGTTATTCTCTGAGAACTCTTGAGCTGAGACCGAGA
9	181	CAATGTTATTCTCTGAGAACTCTTGAGCTGAGACCGAGA
QY	361	GTTCCTCCGGAACTCCAGATTCTCTCTCAANGAGAAAC
QY	361	GTTCCTCCGGAACTCCAGATTCTCTCTCAANGAGAAAC
QY	421	GTCCCGAGCCCTCCAGGCGAACTTAAAGAGCAGGGCT
QY	421	GTCCCGAGCCCTCCAGGCGAACTTAAAGAGCAGGGCT
QY	481	GGCACTGGCTTTAACACCTAGAGCTTCCGGAGTGCTG
QY	481	GGCACTGGCTTTAACACCTAGAGCTTCCGGAGTGCTG
QY	541	GGAAACGGCTGCGACAACTTAAGAGCAGGGCTG
QY	541	GGAAACGGCTGCGACAACTTAAGAGCAGGGCTG
QY	601	GAACTCCACATCATGCTCTTCCTGCCTTCTGGAGG
QY	601	GAACTCCACATCATGCTCTTCCTGCCTTCTGGAGG
Db	660	61 GAACTCCACATCATGCTCTTCCTGCCTTCTGGAGG
Db	660	61 GAACTCCACATCATGCTCTTCCTGCCTTCTGGAGG
Db	660	61 GAACTCCACATCATGCTCTTCCTGCCTTCTGGAGG
Db	720	61 GAACTCCACCTTAATATGCTGTTGCTCCCAA
Db	720	61 GAACTCCACCTTAATATGCTGTTGCTCCCAA
Db	780	721 GTCGCTTTCAGGCCAGTCACCTTCAGGCTTTCAGG
Db	780	721 GTCGCTTTCAGGCCAGTCACCTTCAGGCTTTCAGG
Db	840	781 AGAAGCTGAGAGCTCCAGGCTTCTTAAGAGG
Db	840	781 AGAAGCTGAGAGCTCCAGGCTTCTTAAGAGG
Db	840	781 AGAAGCTGAGAGCTCCAGGCTTCTTAAGAGG
Db	840	841 GAAAGTAGGTCTCTGGTTGATGATCTGGCTG
Db	840	841 GAAAGTAGGTCTCTGGTTGATGATCTGGCTG
Db	900	841 GAAAGTAGGTCTCTGGTTGATGATCTGGCTG
Db	900	841 GAAAGTAGGTCTCTGGTTGATGATCTGGCTG
Db	960	901 GAAAGTAGGTCTCTGGTTGATGATCTGGCTG
Db	1020	951 TCGATGTTGATGATGATGATGATGATGATGATGATG
Db	1020	951 TCGATGTTGATGATGATGATGATGATGATGATGATG
Db	1080	1021 ACAAAACCAAGCTGATGAGACAACTTAATGGGGCT
Db	1080	1021 ACAAAACCAAGCTGATGAGACAACTTAATGGGGCT
Db	1080	1021 ACAAAACCAAGCTGATGAGACAACTTAATGGGGCT
Db	1140	1081 TTGGCAGCTTGAGAGCGCTACAGAGCTGCTGCTG
Db	1140	1081 TTGGCAGCTTGAGAGCGCTACAGAGCTGCTGCTG
Db	1140	1081 TTGGCAGCTTGAGAGCGCTACAGAGCTGCTGCTG
Db	1200	1141 AGCACTGCTTAACTTAAAGCTCTAGAGCAACTCTCT
Db	1200	1141 AGCACTGCTTAACTTAAAGCTCTAGAGCAACTCTCT
Db	1200	1141 AGCACTGCTTAACTTAAAGCTCTAGAGCAACTCTCT
Db	1260	1201 CTGAGATGCTTCCAGCTCTCTCTCTCTCTCTCTCT
Db	1260	1201 CTGAGATGCTTCCAGCTCTCTCTCTCTCTCTCTCT
Db	1320	1261 CCTCTGGCCATCTCTCACGCTATCCAGCTGCTG
Db	1320	1261 CCTCTGGCCATCTCTCACGCTATCCAGCTGCTG
Db	1320	1261 CCTCTGGCCATCTCTCACGCTATCCAGCTGCTG

Gencore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
Run on: August 13, 2003, 04:30:47 ; Search time 5268.5 Seconds 11173.755 Million cell updates/sec				
Title: US-09-818-143-6				
Perfect score: 1439				
Sequence: 1 gggatccctgtttaatca.....agaaaaaaaaaaaaaaaaa 1439				
Scoring table: IDENTITY_NUC				
Gapop 10.0 , Gapext 1.0				
Searched: 288711 seqs, 20454813386 residues				
Post-processing: Minimum Match 0*				
1 number of hits satisfying chosen parameters: 5777422				
Minimum DB seq length: 0				
Maximum DB seq length: 200000000				
Database :				
GenBmlb,*				
1: gb_hab,*				
2: gb_hbg,*				
3: gb_inl,*				
4: gb_oni,*				
5: gb_ort,*				
6: gb_pbt,*				
7: gb_phb,*				
8: gb_pj1,*				
9: gb_ppr,*				
10: gb_ro,*				
11: gb_sts,*				
12: gb_sy,*				
13: gb_un,*				
14: gb_vrl,*				
15: em_bai,*				
16: em_fur,*				
17: em_hum,*				
18: em_in,*				
19: em_mu,*				
20: em_om,*				
21: em_or,*				
22: em_oyr,*				
23: em_pat,*				
24: em_ph,*				
25: em_pl,*				
26: em_rol,*				
27: em_sbs,*				
28: em_un,*				
29: em_vrl,*				
30: em_htg_hum,*				
31: em_htg_inv,*				
32: em_htg_other,*				
33: em_htg_mus,*				
34: em_htg_pln,*				
35: em_htg_rid,*				
36: em_htg_mam,*				
37: em_htg_vrt,*				
38: em_sy,*				
39: em_htgo_hum,*				
40: em_htgo_mus,*				
41: em_htgo_other,*				
RESULT 1				
LOCUS BC03030	BC03030	157 bp	mRNA	linear
DEFINITION Homo sapiens, DKRZP43N161 protein, clone MGC:3507 IMAGE:3605290,				
ACCESSION BC03030			mRNA, complete cds.	
VERSION BC03030.1	GI:30987193			
KEYWORDS MGC.				
SOURCES Homo sapiens (human)				
ORGANISM Homo sapiens				
REFERENCE Makiyama, Metacida; Craniata; Vertebrata; Buteleostomi;				
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE Straubberg, R. Direct Submission				

Pred. No. is the number of results predicted by chance to have a

REMARK
JOURNAL
COMMENT
 USA
 MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgbps-b@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbiology.org>
 Institute, 31 Center Drive, Bethesda, MD 20892-2590,
 Anup Madan, Jessica Pahey, Erin Helton, Mark Kettman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
FEATURES
SOURCE
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>
 Series: ICGU Plate: II Row: P Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7661583.
LOCATION/QUALIFIERS
 1. 1..1557
 /organism "Homo sapiens"
 /mol_type "mRNA"
 /db_xref "LocusID:25939"
 /db_xref "taxon:9606"
 /clone "MGC:3507 IMAGE:3605290"
 /feature_type "uterine_endometrium_adenocarcinoma"
 /clone_id "MIF_MGC_4"
 /lab_host "DHB:R"
 /note "Vector: pORT7"
 275..574
 /codon_start=1
 /product "DKPZP434N161 protein"
 /protein_id "AAH0030.1"
 /db_xref "GI:2098194"
 /translation "MEALLAVNPFRKSFRPDDAKESGSQGFCVPAARPOTMVNGPS
 CSSFLQNLSPQRKRNRAACWQNAQDAPKPNPVCRLKVGRPOASOKLKETGLC"
CDS
ASE COUNT
 RIGIN 441 a 367 c 381 g 368 t
Query Match
 Best Local Similarity 99.4%; Score 1314.8; DB 9; Length 1557;
 Matches 1328; Conservative 0; Mismatches 2; Indels 2; Gaps 1;
V
 108 CCTGTCAGAAGAGAGAGAAAAGAGAGCAATTACCGAACCCCTCCCTCCCCAC 167
 215 CCTGTCAGAAGAGAGAAAAGAGAGCAATTACCGAACCCCTCCCTCCCCAC 274
 168 ATGGAGCTTGGCAAATTTAAATTTCCTAGAAATCTCGAGCTTAAAGAGCGGA 227
 275 ATGGAGCTTGGCAAATTTAAATTTCCTAGAAATCTCGAGCTTAAAGAGCGGA 334
 228 AAGAATCTGGCTCAGAATGGCTCGCGGCCGCGCCGCGCCGCGCCGCGCCG 287
 335 AAAGAACTTGCTCTAGGTTGCTCGGCCGCGCCGCGCCGCGCCGCGCCG 394
 b 288 ACAGGGCCGCTCTGTCCTCCCGGAGCTCCAGAAATTCTCTCTCTCAAGGAGAAC 347
 395 ACAGGGCCGCTCTGTCCTCCCGGAGCTCCAGAAATTCTCTCTCTCAAGGAGAAC 347
 b 408 TTAAGTGTGGTCCAGGCTCTCAAGCCAACTTAAAGAGAGAGCGGCTTGTGA 454
 348 AGGGCATGCGCTTGTGGAAACGCCAGGCGAGGCCGCGCCGCGCCGCGCCG 407
 455 AGGGCATGCGCTTGTGGAAACGCCAGGCGAGGCCGCGCCGCGCCGCGCCG 514
 Y 515 TTAAGTGTGGTCCAGGCTCTCAAGCCAACTTAAAGAGAGAGCGGCTTGTGA 467
 468 AACCAAACATGGCCAGGTGGCTTMAACAATTAGAGACTTCCGGACTGTGCTGG 527
 575 AAACCAAACATGGCCAGGTGGCTTMAACAATTAGAGACTTCCGGACTGTGCTGG 634

RESULT 2
 BC049201
LOCUS BC049201 2017 bp mRNA linear PRI 24-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:39233697, mRNA.
ACCESSION BC049201
VERSION BC049201.1 GI:29165708
KEYWORDS